
Appendix D

Wet Weather Model Configuration, Calibration, and Validation

D Wet Weather Model

Wet weather sources of bacteria are generally associated with wash-off of loads accumulated on the land surface. During rainy periods, these bacteria loads are delivered to the waterbody through creeks and stormwater collection systems. Often, bacteria sources can be linked to specific land use types that have higher relative accumulation rates of bacteria, or are more likely to deliver bacteria to waterbodies due to delivery through stormwater collection systems. To assess the link between sources of bacteria and the impaired waters, a modeling system may be utilized that simulates the build-up and wash-off of bacteria and the hydrologic and hydraulic processes that affect delivery. Understanding and modeling of these processes provides the necessary decision support for TMDL development and allocation of loads to sources.

The wet weather TMDL calculation was based on a watershed model of the drainage area associated with each impaired waterbody. USEPA's Loading Simulation Program in C++ (LSPC) was selected to simulate the hydrologic processes and bacteria loading to receiving waterbodies in the San Diego Region. LSPC is a component of the EPA's TMDL Modeling Toolbox (Toolbox), which has been developed through a joint effort between USEPA and Tetra Tech, Inc. It integrates a geographical information system (GIS), comprehensive data storage and management capabilities, a dynamic watershed model (a re-coded version of EPA's Hydrological Simulation Program – FORTRAN [HSPF]), and a data analysis/post-processing system into a convenient PC-based windows interface that dictates no software requirements.

An LSPC model was configured for many of the watersheds in the San Diego Region and was then used to simulate series of hydraulically connected subwatersheds. Configuration of the model involved subdividing the watersheds within the San Diego Region into modeling units, followed by continuous simulation of flow and water quality for those units using meteorological, land use, soils, stream, point source, and bacteria representation data. Development and application of the watershed model to address the project objectives involved a number of important steps:

1. Watershed Segmentation
2. Configuration of Key Model Components
3. Model Calibration and Validation

D.1 Watershed Segmentation

Watershed segmentation refers to the subdivision of all watersheds in the San Diego Region into smaller, discrete subwatersheds for modeling and analysis. This subdivision was primarily based on the stream networks and topographic variability, and secondarily on the locations of flow and water quality monitoring stations, consistency of hydrologic factors, land use consistency, and existing watershed boundaries (based on CALWTR 2.2 watershed boundaries). The San Diego Region was divided into sixteen basins for model

configuration and subwatershed delineation—thirteen basins were modeled for assessment of bacteria loads to impaired waterbodies; three additional watersheds (Santa Margarita River, Tecolote Creek, and Rose Creek) were configured for region-wide calibration, since data in these watersheds were plentiful. Basins and respective subwatershed delineations are presented in Appendix C.

D.2 Configuration of Key Model Components

Configuration of the watershed model involved consideration of four major components: meteorological data, land use representation, hydrologic and pollutant representation, and waterbody representation. These components provided the basis for the model's ability to estimate flow and pollutant loadings. Meteorological data essentially drive the watershed model. Rainfall and other parameters are key inputs to LSPC's hydrologic algorithms. The land use representation provides the basis for distributing soils and pollutant loading characteristics throughout the basin. Hydrologic and pollutant representation refers to the LSPC modules or algorithms used to simulate hydrologic processes (e.g., surface runoff, evapotranspiration, and infiltration) and pollutant loading processes (primarily accumulation and washoff). Waterbody representation refers to LSPC modules or algorithms used to simulate flow and pollutant transport through streams and rivers.

D.2.1 Meteorology

Meteorological data are a critical component of the watershed model. LSPC requires appropriate representation of precipitation and potential evapotranspiration. In general, hourly precipitation (or finer resolution) data are recommended for nonpoint source modeling. Therefore, only weather stations with hourly-recorded data were considered in the precipitation data selection process. Rainfall-runoff processes for each subwatershed were driven by precipitation data from the most representative station. These data provide necessary input to LSPC algorithms for hydrologic and water quality representation.

Meteorological data have been accessed from a number of sources in an effort to develop the most representative dataset for the San Diego Region. Hourly rainfall data were obtained from the National Climatic Data Center (NCDC) of the National Oceanic and Atmospheric Administration (NOAA), the Automatic Local Evaluation in Real Time (ALERT) Flood Warning System managed by the County of San Diego, and the California Irrigation Management Information System (CIMIS). The above data were reviewed based on geographic location, period of record, and missing data to determine the most appropriate meteorological stations. Ultimately, meteorological data were utilized from 16 area weather stations for January 1990-September 2002 (Figure D-1).

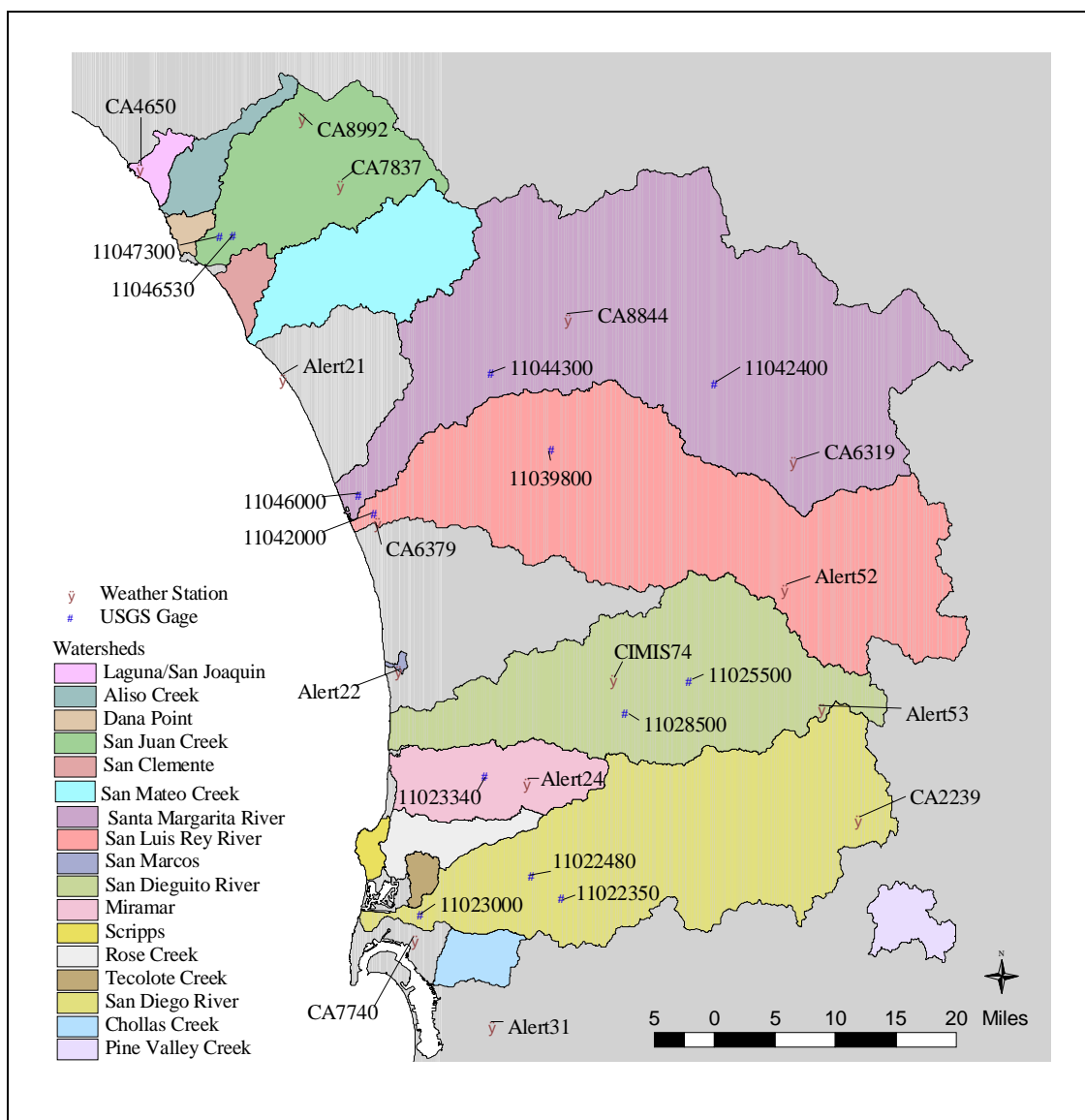


Figure D-1. Weather stations and flow gages utilized for wet weather modeling

Long-term hourly wind speed, cloud cover, temperature, and dew point data are available for a number of weather stations in the San Diego Region. Data from Lindbergh Field, the San Diego Airport (COOP ID #047740), were obtained from NCDC for characterization of meteorology of the modeled watersheds. Using this data, the METCMP utility, available from USGS, was used to calculate hourly potential evapotranspiration.

D.2.2 Land Use Representation

The watershed model requires a basis for distributing hydrologic and pollutant loading parameters. This is necessary to appropriately represent hydrologic variability throughout the basin, which is influenced by land surface and subsurface characteristics.

It is also necessary to represent variability in pollutant loading, which is highly correlated to land practices. The basis for this distribution was provided by land use coverage of the entire watershed.

Three sources of land use data were used in this modeling effort. The primary source of data was the San Diego Association of Governments (SANDAG) 2000 land use dataset that covers San Diego County. This dataset was supplemented with land use data from the Southern California Association of Governments (SCAG) for Orange County and portions of Riverside County. A small area in Riverside County was not covered by either land use dataset. To obtain complete coverage, the 1993 USGS Multi-Resolution Land Characteristic data was used to fill this remaining data gap.

Although the multiple categories in the land use coverage provide much detail regarding spatial representation of land practices in the watershed, such resolution is unnecessary for watershed modeling if many of the categories share hydrologic or pollutant loading characteristics. Therefore, many land use categories were grouped into similar classifications, resulting in a subset of 13 categories for modeling. Selection of these land use categories was based on the availability of monitoring data and literature values that could be used to characterize individual land use contributions and critical bacteria-contributing practices associated with different land uses. For example, multiple urban categories were represented independently (e.g., high density residential, low density residential, and commercial/institutional), whereas forest and other natural categories were grouped. Table D-1 presents the land use distribution in each of the thirteen watersheds contributing to waterbody impairments.

LSPC algorithms require that land use categories be divided into separate pervious and impervious land units for modeling. This division was made for the appropriate land uses (primarily urban) to represent impervious and pervious areas separately. The division was based on typical impervious percentages associated with different land use types from the Soil Conservation Service's TR-55 Manual (Soil Conservation Service, 1986).

Table D-1. Land use Areas (square miles) of Each Impaired Watershed

Watershed	Low Density Residential (1100)	High Density Residential (1200)	Commercial/Institutional (1400)	Industrial/Transportation (1500)	Military (1600)	Parks/Recreation (1700)	Open Recreation (1800)	Agriculture (2000)	Dairy/Intensive Livestock (2400)	Horse Ranches (2700)	Open Space (4000)	Water (5000)	Transitional (7000)	Total
Laguna / San Joaquin	2.39	0.61	0.34	0.11	0.00	0.18	0.02	0.00	0.00	0.02	10.02	0.00	0.23	13.94
Aliso Creek	8.75	3.76	2.14	0.89	0.00	0.69	0.40	0.07	0.00	0.03	16.09	0.06	2.86	35.74
Dana Point	3.51	1.30	0.25	0.01	0.00	0.28	0.32	0.00	0.00	0.00	2.70	0.00	0.53	8.89
San Juan Creek	15.61	2.97	3.09	2.90	0.00	1.03	1.86	7.57	0.00	0.40	137.07	0.66	4.03	177.18
San Clemente	3.85	1.31	0.66	1.17	0.02	0.37	0.52	0.00	0.00	0.00	10.06	0.00	0.81	18.78
San Luis Rey River	42.86	4.22	3.24	4.92	15.31	1.65	2.56	123.49	8.51	0.00	350.46	2.56	0.63	560.42
San Marcos	0.34	0.17	0.19	0.05	0.00	0.04	0.10	0.06	0.25	0.00	0.13	0.01	0.10	1.43
San Dieguito River	43.58	2.26	5.33	2.22	0.00	1.19	3.19	61.72	5.71	0.00	215.96	2.72	2.34	346.22
Miramar	22.42	3.86	11.41	3.28	0.00	1.70	1.14	2.29	0.93	0.00	44.47	0.26	1.96	93.73
Scripps	5.21	1.32	0.86	0.05	0.00	0.13	0.20	0.00	0.00	0.00	0.94	0.01	0.03	8.75
San Diego River	65.65	10.61	16.36	10.07	3.07	2.73	2.06	9.46	0.87	0.00	308.67	6.44	0.50	436.48
Chollas Creek	14.75	2.87	3.79	1.61	0.02	0.38	0.52	0.00	0.00	0.00	2.73	0.03	0.09	26.80
Pine Valley Creek	0.13	0.00	0.03	0.00	0.00	0.11	0.00	0.03	0.00	0.00	29.10	0.13	0.00	29.53

D.2.3 Hydrology Representation

The LSPC PWATER (water budget simulation for pervious land segments) and IWATER (water budget simulation for impervious land segments) modules, which are identical to those in HSPF, were used to represent hydrology for all pervious and impervious land units (Bicknell et al., 1996). Designation of key hydrologic parameters in the PWATER and IWATER modules of LSPC were required. These parameters are associated with infiltration, groundwater flow, and overland flow. USDA's STATSGO Soils Database served as a starting point for designation of infiltration and groundwater flow parameters. For parameter values not easily derived from these sources, documentation on past HSPF applications were accessed, particularly the recent modeling studies performed for the San Jacinto River Watershed (Tetra Tech, Inc, 2003) and Santa Monica Bay (LARWQCB, 2002). Starting values were refined through the hydrologic calibration process (described in the next section).

D.2.4 Pollutant Representation

Loading processes for FC, TC, and ENT were represented for each land unit using the LSPC PQUAL (simulation of quality constituents for pervious land segments) and IQUAL (simulation of quality constituents for impervious land segments) modules, which are identical to those in HSPF. These modules simulate the accumulation of pollutants during dry periods and the washoff of pollutants during storm events. Starting values for parameters relating to land-use-specific accumulation rates and buildup limits, were obtained from a study performed by the Southern California Coastal Water Research Project (SCCWRP) to support bacteria TMDL development of Santa Monica Bay (LARWQCB, 2002). These starting values served as baseline conditions for water quality calibration; the appropriateness of these values to the San Diego Region watershed was validated through comparison to local water quality data. Although atmospheric deposition may be an issue in the watersheds, it was not explicitly simulated in the watershed model. It was, however, represented implicitly in the model through use of the land use- and pollutant-specific accumulation rates.

There were six major inland dischargers during the simulation period and these were incorporated into the LSPC model as point sources of flow and bacteria. Each point source is located in the Santa Margarita River watershed – five at Camp Pendleton and one along Murrieta Creek (Santa Rosa Water Reclamation Facility). Although the Santa Margarita River watershed had no waterbodies impaired for bacteria, it was simulated in this wet weather modeling effort due to the availability of streamflow and bacteria monitoring data, which were used for hydrologic and water quality calibration and validation. It is important to note that all six major inland discharges were eliminated by 2002.

D.2.5 Waterbody Representation

Each delineated subwatershed was represented with a single stream assumed to be completely mixed, one-dimensional segments with a trapezoidal cross-section. The

National Hydrography Dataset (NHD) stream reach network for USGS hydrologic units 18070301 through 18070305 were used to determine the representative stream reach for each subwatershed. Once the representative reach was identified, slopes were calculated based on DEM data and stream lengths measured from the original NHD stream coverage. In addition to stream slope and length, mean depths and channel widths are required to route flow and pollutants through the hydrologically connected subwatersheds. Mean stream depth and channel width were estimated using regression curves that relate upstream drainage area to stream dimensions. An estimated Manning's roughness coefficient of 0.2 was also applied to each representative stream reach.

In addition to the streams which route flow and transport pollutants through the watersheds, there were several reservoirs within the region that were large enough to impound a significant portion of flow during wet periods. To represent these reservoirs in the watershed model, the length, width, maximum depth, infiltration rate, and spillway height and width were obtained for each reservoir. The reservoirs impounded all upstream flow until the water depth exceeded the spillway height, causing overflow and thus contributing to downstream flow and bacteria loading.

D.3 Model Calibration and Validation

After the model was configured, model calibration and validation were performed. This is generally a two-phase process, with hydrology calibration and validation completed before repeating the process for water quality. Upon completion of the calibration and validation at selected locations, a calibrated dataset containing parameter values for each modeled land use and pollutant was developed.

Calibration refers to the adjustment or fine-tuning of modeling parameters to reproduce observations. The calibration was performed for different LSPC modules at multiple locations throughout the watershed. This approach ensured that heterogeneities were accurately represented. Subsequently, model validation was performed to test the calibrated parameters at different locations or for different time periods, without further adjustment. To ensure that the model results are as current as possible and to provide for a range of hydrologic conditions, January 1991 through September 2002 was selected as the time period for simulation.

D.3.1 Hydrology Calibration and Validation

Hydrology is the first model component calibrated because estimation of bacteria loading relies heavily on flow prediction. The hydrology calibration involves a comparison of model results to in-stream flow observations at selected locations. After comparing the results, key hydrologic parameters were adjusted and additional model simulations were performed. This iterative process was repeated until the simulated results closely represented the system and reproduced observed flow patterns and magnitudes.

Gaging stations representing diverse hydrologic regions of the San Diego Region were used for calibration, including eleven USGS streamflow gage stations (Table D-2 and Figure D-1). These gaging stations were selected because they either had a robust historical record or they were in a strategic location (i.e. along a 303(d) listed waterbody, downstream of a reservoir, or along an otherwise unmonitored reach).

Table D-2. USGS Station Used For Hydrology Calibration and Validation

Station Number	Station Name	Historical Record	Selected Calibration Period	Selected Validation Period	Watershed and Model Subwatershed
11022480	San Diego River at Mast Road near Santee, CA	5/1/1912 - 9/30/2002	1/1/1991 - 12/31/1996	1/1/1997 - 12/31/2001	San Diego River (1805)
11023000	San Diego River at Fashion Valley at San Diego, CA	1/18/1982 - 9/30/2002	1/1/1991 - 12/31/1996	1/1/1997 - 12/31/2001	San Diego River (1801)
11023340	Los Penasquitos Creek near Poway, CA	10/1/1964 - 9/30/2002	1/1/1991 - 12/31/1996	1/1/1997 - 12/31/2001	Miramar (1406)
11025500	Santa Ysabel Creek near Ramona, CA	2/1/1912 - 9/30/2002	1/1/1991 - 12/31/1996	1/1/1997 - 12/31/2001	San Dieguito (1316)
11028500	Santa Maria Creek near Ramona, CA	12/1/1912 - 9/30/2002	1/1/1991 - 12/31/1996	1/1/1997 - 12/31/2001	San Dieguito (1324)
11042000	San Luis Rey River at Oceanside, CA	10/1/1912 - 11/10/1997; 4/29/1998 - 9/30/2002	9/1/1993 - 8/31/1997	5/1/1998 - 4/30/2002	San Luis Rey (702)
11042400	Temecula Creek near Aguanga, CA	8/1/1957 - 9/30/2002	1/1/1991 - 12/31/1996	1/1/1997 - 12/31/2001	Santa Margarita (658)
11044300	Santa Margarita River at FPU D Sump near Fallbrook, CA	10/1/1989 - 9/30/2002	1/1/1991 - 12/31/1996	1/1/1997 - 12/31/2001	Santa Margarita (615)
11046000	Santa Margarita River at Ysidora, CA	3/1/1923 - 2/25/1999; 10/1/2001 - 9/30/2002	1/1/1991 - 12/31/1995	1/1/1996 - 12/31/1998	Santa Margarita (602)
11046530	San Juan Creek at La Novia Street Bridge near San Juan Capistrano, CA	10/1/1985 - 9/30/2002	1/1/1991 - 12/31/1996	1/1/1997 - 12/31/2001	San Juan (411)
11047300	Arroyo Trabuco near San Juan Capistrano, CA	10/1/1970 - 9/30/1989; 10/1/1995 - 9/30/2002	10/1/1995 - 4/30/1999	5/1/1999 - 4/30/2002	San Juan (403)
11022350	Forester Creek near El Cajon, CA	10/1/1993 - 9/30/2002	none (insufficient period of record)	1/1/1991 - 9/30/1993	San Diego River (1843)
11039800	San Luis Rey River at Couser Canyon Bridge near Pala, CA	10/1/1986 - 1/4/1993	none (insufficient period of record)	1/1/1991 - 12/31/1992	San Luis Rey (711)

The calibration years were selected based on annual precipitation variability and the availability of observation data to represent a continuum of hydrologic conditions: low, mean, and high flow. Calibration for these conditions was necessary to ensure that the model would accurately predict a range of conditions over a longer period of time.

Key considerations in the hydrology calibration included the overall water balance, the high-flow/low-flow distribution, stormflows, and seasonal variation. At least two criteria

for goodness of fit were used for calibration: graphical comparison and the relative error method. Graphical comparisons were extremely useful for judging the results of model calibration; time-variable plots of observed versus modeled flow provided insight into the model's representation of storm hydrographs, baseflow recession, time distributions, and other pertinent factors often overlooked by statistical comparisons. The model's accuracy was primarily assessed through interpretation of the time-variable plots. The relative error method was used to support the goodness of fit evaluation through a quantitative comparison.

After calibrating hydrology at the eleven locations, a validation of these hydrologic parameters was made through a comparison of model output to different time periods at the same gages as well as two additional gages (Table D-2). The validation essentially confirmed the applicability of the regional hydrologic parameters derived during the calibration process. Validation results were assessed in a similar manner to calibration: graphical comparison and the relative error method.

Hydrology calibration and validation results, including time series plots and relative error tables, are presented for each gage in Appendix E. The calibration results, which are presented first, include graphs to represent overall model fit, seasonal trends, and two time series plots. These graphs are followed by a table that quantifies the model results and observed gage data. This table also provides relative errors between the modeled and observed values in the storm volumes and highest flows. The presentation of model validation results follows the calibration tables and graphs for each gage. Two additional gages that had a limited historical record were used as additional validation. Validation was assessed through a time series plot and a relative error table identical to the calibration table.

Overall, during model calibration the model predicted storm volumes and storm peaks well. Since the runoff and resulting streamflow is highly dependent on rainfall, occasional storms were over-predicted or under-predicted depending on the spatial variability of the meteorologic and gage stations. The validation results also showed a good fit between modeled and observed values, thus confirming the applicability of the calibrated hydrologic parameters to the San Diego Region.

D.3.2 Water Quality

After the model was calibrated and validated for hydrology, water quality simulations were performed. As described above, previously calibrated, land use specific accumulation and maximum build up rates for fecal coliforms, total coliforms, and enterococci (LARWQCB, 2002) were used for the water quality simulations. Since these values have been successfully applied to recent bacteria models, including TMDLs, in southern California, they were considered to be sufficiently calibrated. Therefore, the water quality simulations were used to further validate these rates. The objective of the validation process was to best represent bacteria concentrations during storm events at monitoring stations throughout the region.

Only data from wet weather events (rainfall of 0.2 inches or greater and the following 72 hours) were used for comparison with model water quality output. This greatly reduced the availability of bacteria monitoring data for use in the validation process; however, it was important to differentiate between wet and dry periods due to the separate approaches utilized for this TMDL. There were 107 monitoring stations in the modeled subwatersheds with wet weather monitoring data that overlapped with the modeling period (Tables D-3 through D-5). The spatial variability of these locations was excellent (ranging from urban to open land uses); however, the temporal variability and total number of samples limited statistical analysis to basinwide summary statistics rather than comprehensive time series and relative error analyses at each monitoring location.

Table D-3. Basinwide water quality data used for fecal coliform validation

Basin	Number of		Fecal Coliform (MPN/100mL)		
	Sites	Samples	Minimum	Mean	Maximum
Aliso Creek	59	217	2	11,142	160,000
San Juan Creek	7	9	200	4,222	26,000
Santa Margarita River	14	83	2	1,204	50,000
Rose Creek & Tecolote Creek	17	30	31	9,939	137,400
San Diego River	6	36	2	1,557	24,000

Table D-4. Basinwide water quality data used for total coliform validation

Basin	Number of		Total Coliform (MPN/100mL)		
	Sites	Samples	Minimum	Mean	Maximum
Aliso Creek	56	206	2	32,246	160,000
San Juan Creek	7	9	680	16,356	70,000
Santa Margarita River	14	36	230	3,248	50,000
Rose Creek & Tecolote Creek	15	24	4,884	333,384	2,419,200
San Diego River	6	34	300	14,885	300,000

Table D-5. Basinwide water quality data used for enterococcus validation

Basin	Number of		Enterococcus (MPN/100mL)		
	Sites	Samples	Minimum	Mean	Maximum
Aliso Creek	59	217	1	3,720	72,000
San Juan Creek	7	9	340	8,056	51,000
Rose Creek & Tecolote Creek	17	29	20	6,978	32,550
Pine Valley Creek	4	24	1	1,065	20,000

To assess model fit with available data, the time series model output was graphically compared to the observed data. Appendix F (Figures 1-11) presents time series graphs of modeled and observed data for downstream subwatersheds with a reasonable number of samples. Ensuring that the storm events were represented within the range of the data

over time is the most practical and meaningful means of assessing the quality of the model output. The time series plots indicate that the model predicts the FC, TC, and ENT concentrations within the range of observed data (ranges of observed data are presented in Tables D-3 through D-5) and at a similar frequency. This is especially evident in subwatersheds where there is a significant amount of data across a wide temporal range (see Appendix F, Figure 1).

To provide a side-by-side comparison of the available wet weather monitoring data with model output for the same day, data were grouped by basin to increase sample size. Graphs of concentration by percentile of unit area flow (inches/acre) are presented in Appendix F (Figures 12-25) for each pollutant in the basins where data were available. Presenting the data as a function of flow facilitates analysis of the results which are pertinent to the wet weather model. Specifically, the higher flows (larger percentiles) are likely associated with the actual precipitation event, rather than the assumed wet period of 72 hours following the storm. For lower flows, observed data that met the wet weather criterion (0.2 inches of rainfall and following 72 hours) may not be representative of true wet conditions, which explains the deviance between model predictions and ranges of observed water quality. However, dry periods are addressed in a separate approach in this TMDL with better accuracy.

Figures 12 through 25 in Appendix F depict the average and range for observed and modeled FC, TC, and ENT concentrations in the basins identified above. These graphs indicate that the model compared well to observed data, especially for basins with larger sample sizes and in the larger unit area flow percentiles. Discrepancies may be due to small sample sizes, the variability in bacteria monitoring and analysis, or the range of time defined as a wet period (72 hours after a 0.2 inch or greater storm).

Analysis of the time series graphs and the unit area flow summary plots indicate that the previously calibrated bacteria accumulation and maximum build-up rates (LARWQCB, 2002) are applicable, and therefore validated, for the San Diego region. Additional bacteriological data collection is likely to further support these findings considering that the model matched observed data extremely well for all three pollutants when an abundance of observed wet weather data was available (see Appendix F, Figures 12-14).

D.4 Application of Wet Weather Model

After completing model calibration and validation for hydrology and water quality, the model was applied to obtain hourly output for the critical period described in Section 6.1.1. The maximum hourly FC, TC, and ENT concentrations were obtained for each wet day in the critical period (1993) for all subwatersheds associated with a 303(d) listed segment. These concentrations, along with their associated average daily flow, were used to generate TMDL load duration curves (Appendices G and H). The overall load capacity was incorporated into the load duration curves. Predicted loads that fell above the load capacity are exceedances and were then divided by the total existing load to calculate the percent reduction required to achieve the beneficial use of the receiving waterbody.